

Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/141552/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Jones, Raya A. ORCID: <https://orcid.org/0000-0002-5419-677X> 2022.
Projective anthropomorphism as a dialogue with ourselves. International Journal of Social Robotics 14 , pp. 2063-2069. 10.1007/s12369-021-00793-7
file

Publishers page: <https://doi.org/10.1007/s12369-021-00793-7>
<<https://doi.org/10.1007/s12369-021-00793-7>>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies.

See

<http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.





Projective Anthropomorphism as a Dialogue with Ourselves

Raya A. Jones¹

Accepted: 27 April 2021
© The Author(s) 2021

Abstract

This paper introduces an original concept (projective anthropomorphism) towards exploring a psychological dimension that is irreducible to the forms of anthropomorphism investigated in both cognitive science and social robotics. Projective anthropomorphism is an unconscious bias towards anticipating humanlike characteristics in robots. An overview of the variety of ways in which projection has been conceptualised in psychology and psychoanalysis is provided before discussing implications for theorising projective anthropomorphism. The proposed concept alludes to the projection of existential anxieties and desires onto myths, legends, linguistic tropes, and science-fiction motifs of humanoid automata. Such motifs and their associated narratives populate contemporary popular culture, and feed into social representations of robots. The importance of considering projective anthropomorphism lies in the extent to which its phenomena channel people's expectations and attitudes towards technological artefacts, as well as steering technological possibilities.

Keywords Anthropomorphism · Psychological projection · Human–robot interaction

1 Introduction

Investigations and discussions of anthropomorphism in human–robot interaction (HRI) settings typically pertain to persons' interactions with an actual machine, whether the robot is physically present, an image on a screen, or a design concept. Anthropomorphic phenomena in this context vary along a dimension from involuntary reactions to a robot's physical appearance or behaviour as if it were human (reactive anthropomorphism) to deliberately designing robots with humanlike features (proactive anthropomorphism). The heterogeneity of such phenomena means that various manifestations are likely to differ in terms of their underlying psychological mechanisms, triggers, motivations, and functions.

Moreover, a robot is not only a machine that engineers build but is also a semiotic object, comprised of signs and symbols, and constructed in discourse and imagination [1]. Images of manmade humanlike automata have populated the imagination for millennia in myths, legends, and modern popular culture. From the standpoint of depth psychology

(psychoanalytical and other psychodynamic perspectives), these motifs could be understood as products of a symbol-forming mechanism whereby unconscious preoccupations are brought into conscious awareness through projections in the form of concrete images, tales, and artefacts. In this sense, anthropomorphic motifs are emergent properties of a tacit dialogue with our 'inner' selves, collectively as well as individually.

Projective anthropomorphism (a term coined here) concerns this aspect of the discourse of social robotics. It is definable as an unconscious bias towards anticipating humanlike characteristics in robots. Projective anthropomorphism is irreducible to either reactive or proactive forms of anthropomorphism, as argued in Sects. 2 and 3. Put another way, projective anthropomorphism cannot be placed on the bipolar reactive-proactive dimension, but could be construed as an orthogonal dimension that intersects with the former, and potentially complements our understanding of robotics as a human endeavour in general. Section 4 reviews psychological concepts of projection so as to identify the present application of the term. Having set conceptual parameters, Sect. 5 provides some examples. Finally, Sect. 6 returns to the idea that a robot is also a semiotic object [1]. Drawing upon psychoanalytical theories, this section speculatively considers psychological functions that the compulsion to imagine humanoid automata might serve.

✉ Raya A. Jones
JonesRA9@cardiff.ac.uk

¹ School of Social Sciences, Cardiff University, Glamorgan Building, King Edward VII Avenue, Cardiff CF10 3WT, UK

2 Explaining Anthropomorphic Reactions

The extant literature on the psychological phenomenon of anthropomorphism centres on describing and explaining people's anthropomorphic reactions. It is generally believed that the activation of anthropomorphic stereotypes in people through robots' humanlike appearance or behaviour can bring a sense of familiarity, confidence and simplicity into people's interactions with these novel agents. This in turn may encourage trust and acceptance of social robots. Robot design and marketing benefits from fact-finding empiricist research that identifies conditions under which people are more likely to respond to a robot as if it were human, and the extent to which their experience is likely to be associated with positive or negative affect.

Research in cognitive science, on the other hand, seeks to establish universal causal mechanisms that may account for the occurrence of anthropomorphism whenever it happens. The three-factors theory of anthropomorphism outlines three psychological determinants of this phenomenon: (1) the extent to which anthropocentric knowledge is accessible and applicable (elicited agent knowledge); (2) a motivation to explain and understand other agents' behaviour (effectance motivation); and (3) a desire for social contact and affiliation (sociality motivation) [2]. While the theory does not focus solely on robots, its authors point out practical implications for HRI: humanoid features can enhance people's sense of familiarity and affinity when interacting with the robot. Findings reported in [3], for instance, lend support to the first determinant outlined in the three-factors theory, namely that people generalise initial attributions across agent categories.

The authors of [4] similarly postulate a cognitive mechanism characterised by a dual process: initially a fast intuitive process classifies an object as humanlike (implicit anthropomorphism), and is followed with a conscious reflective process that may moderate the initial judgment (explicit anthropomorphism). While empirical support for that hypothesis is tentative at present, its gist has intuitive appeal. Turkle's anecdotal account of her response to Cog when visiting Rodney Brooks' lab at MIT captures this duality: 'My response was involuntary ... visceral. Cog had a face, it made eye contact, and it followed my movements. ... although I knew Cog to be a machine, I had to fight my instinct to react to "him" as a person' [5, p. 84].

Cognitive models such as cited above aim to explain reactions to objects that are available to sensory perception in the here-and-now. Theories in this vein may shed light on what causes anthropomorphic reactions in brief exposures to robots within experimental procedures, as if in a vacuum. Idiographic factors, such as the meaning that the encounter with Cog had for Turkle, fall outside the remit

of the 'classic' scientific paradigm. The visceral experience she described did not remain contained within that moment in the lab. Her experience became interwoven into her narrative, including her reasons for telling about it in her book. Moreover, theories such as the aforementioned do not concern motivations to create humanoid robots, let alone why human beings have been telling imaginative stories of humanoid automata across cultures and millennia.

3 Building Humanlike Robots

From the standpoint of functional design, when the purpose is to build socially interactive robots for everyday applications, the objective is to create a robot that outwardly appears to be socially intelligent though its internal workings may not resemble the human mind [6].

Socially assistive robots (SAR) are sometimes given a humanoid appearance so as to enhance users' sense of familiarity and affinity. For example, Dautenhahn and her team at the University of Hertfordshire have developed Kaspar, a talking robot with a simplified human face and moveable limbs and features [7]. This robot is designed to help children with autism or communication difficulties to develop essential social skills through games such as peekaboo, imitation games, and learning activities. Having the doll-like robot looking like a boy presumably enhances the child's sense of affinity. Other designers of robots for children opt for more cartoon-like or abstract forms. Leka, also developed for work with autistic children, is a small sphere with a minimalist face. Its creator is quoted as saying that when children performing a task with Leka get it right, 'the robot reacts emotionally,' by flashing happy facial expressions; "The kids love it." [8, p. 1]. Giving humanlike characteristics to SARs is conducive to making users experience the interaction as more natural. This strategy responds to the engineering challenges that [1] has termed 'the classic problem of natural interaction in HRI; namely, how to design robots that have at least some human social characteristics' (p. 93). Engineers may harness knowledge of cognitive mechanisms underlying anthropomorphic reactions towards optimising users' positive experience of the interaction. As the example of Leka suggests, humanlike behaviour, speech and emotional expression might be more crucial for people's perception of robot sociality than a humanoid appearance would be.

Engineers who opt for a functional design often give their robots behavioural characteristics of intentional agents though they do not need to understand how human cognition 'really works' [6, p. 147]. In contrast, biologically inspired design aims to create robots that internally mimic human cognition. For example, Cog was originally built by Brooks and his team at MIT in the 1990s. According to the project's

website, the motivation behind creating Cog stemmed from the hypothesis that ‘humanoid intelligence requires humanoid interactions with the world’ [9]. Brooks and cowriters identified two goals of building humanoid robots: ‘an engineering goal of building a prototype general purpose flexible and dextrous autonomous robot and a scientific goal of understanding human cognition’ [10, p. 52]. The scientific goal accords with the synthetic approach in cognitive science, which first became possible with the advent of computers in the late 1940s. It complements the analytic approach, whereby scientists try to understand the mind through systematic observations and experiments. Scientists who opt for the synthetic approach try to understand the mind through modelling cognitive architectures in computers and robots. Within this specialised context, the authors of [10] argue against seeking to understand human cognition by virtue of disembodied computer simulation, since our cognition is essentially embodied (hence the motivation to build Cog).

Whether computers or robots are used for modelling mental processes, however, the shared epistemological position is that scientists can have a candidate explanation for some phenomenon by building an artificial system based on the principles that are believed to underlie the real phenomenon. The possibility of reverse engineering the human mind goes hand in hand with reverse anthropomorphism; that is, looking at ourselves and seeing a machine. As Brooks put it, ‘Humans, after all, are machines made up of organic molecules whose interactions can all be aped (we think) by sufficiently powerful computers’ [11, p. 86]. The entrenched view of man as a machine has been perpetuated since at least 1747, when La Mettrie first published his thesis *L’homme machine*. It has made possible a wide range of technological advances, but it resonates with technological determinism; and, in turn, manifests also in a tendency to reduce the concept of a social interaction to superficial mechanistic aspects, such as taking turns in a conversation or reciprocating emotional expressions [1]. Faith in the possibility of reverse engineering the mind might lead to an expectation that a biologically inspired approach to robot design could result in making a good-enough fake that is closer to having genuine intentionality and self-consciousness. Yet, modelling the human mind on a machine leaves out the very ‘spark’ that motivates humans to contemplate their own nature, and fires some scientists’ ambition to seek this understanding by building humanlike robots.

4 Varieties of Psychological Projection

It could be argued that any form of anthropomorphism involves projection (we project humanlike attributes onto something that does not have them). However, both the words ‘projection’ and ‘projective’ are used in psychology

in a variety of ways, some of which contradict each other insofar as the same terminology refers to fundamentally different psychological functions. This section provides an overview of the main usages before discussing some of the implications for theorising ‘projective anthropomorphism’.

Centring on social cognition, Baumeister and his cowriters [12] draw a distinction between projection as a cognitive bias, on the one side, and projection as a defence mechanism in the Freudian sense, on the other. Conceptualised as a cognitive bias, projection ranges from naively perceiving one’s own traits (good and bad) in other people to inaccurately perceiving others as having traits that one believes one does not have. In contradistinction, Freud conceptualised projection as a defence mechanism whereby perceiving an undesirable trait in others serves to avoid recognising this trait in oneself. Baumeister et al. contend that the cognitive-bias option is well supported with research evidence whereas the Freudian theory has little empirical support. Either way, the multiple ways of understanding projection ‘vary mainly along the dimension of how effectively the undesirable trait or motive is repudiated as part of the self’ [12, p. 1090]. This dimension, however, cannot accommodate the Jungian conceptualisation of projection, for instance, or the so-called Projective Hypothesis that informs certain methods of personality assessment (described later in this section).

In contrast to Freud’s assumption of a mechanism serving ‘to fend off an idea that was intolerable to [the patient’s] ego by projecting its subject-matter into the external world’ (Freud 1895, quoted in [13, p. 2]), C. G. Jung postulated a mechanism of projection as necessary for the integration of the personality (individuation). The Jungian model describes the psyche as an equilibrium-seeking system that comprises conscious and unconscious opposites, such as ego and shadow, gender and counter-gender (anima is the feminine in a man, animus is the masculine in a woman). An unconscious element can be brought to consciousness only when it is seen in other people and is subsequently recognised as a projection (rather than naively believing it to be a trait of the other person): ‘The shadow can be realized only through a relation to a partner, and anima and animus only through a relation to a partner of the opposite sex’ ([14] para. 42). Hence, contrary to the Freudian premise that projection hides some inconvenient truth about oneself, the Jungian construal points to a non-pathological process that functions towards opening up one’s awareness of one’s whole personality.

Freud himself talked about projection also in a more general sense, and his followers further elaborated diverse meanings. As [13] outlines, some psychoanalysts placed the emphasis on projection as a defence mechanism, but the term has acquired a variety of meanings within that context. Another direction was taken by Melanie Klein and her followers, who placed the emphasis on ‘projection as the

process whereby the ego expelled (projected) its own sadistic impulses into the external world' [13, p. 4]. The concept of projective identification (articulated by Klein in 1946) posits a pathological process whereby 'parts of the self and internal objects are split off and projected into the external object, which then becomes possessed by, controlled and identified with the projected parts' (quoted in [15, p. 13]).

The adjective 'projective' in Klein's concept of projective identification indicates a markedly different psychological phenomenon than does the same adjective in the case of projective methods of personality assessment, such as the Rorschach inkblot test and Thematic Apperception Test (TAT). The designation 'protective' was introduced by Frank in 1939 [16] in order to differentiate this methodology from so-called 'objective' methods such as standardised self-report inventories. A projective method entails presenting material or a task that is designed or chosen for pre-empting any preconceived expectations about the desirability of the response. While the Rorschach and TAT require the subject—who may be a client in clinical assessment or a research participant—to respond to a standard set of visual stimuli, the House–Tree–Person (H–T–P) technique requires the subject to make a freehand drawing of a house, tree, and person [17]. Other techniques require sentence completion or manipulating objects (see a review in [18]). An obvious though nontrivial difference between the kind of projection that takes place during the performance of projective tests and the kinds of projection discussed earlier lies in whether the target of projection is a real person or not. People do not 'identify' with an inkblot, for instance, cannot inaccurately perceive the traits of a human figure painted in a TAT picture or 'get wrong' the person they draw in the H–T–P test.

Historically, projective methodology was loosely associated with the psychoanalytical ethos of the early twentieth century, but the multiple techniques that have proliferated since then have no specific theory in common. Epistemologically, these share little more than the so-called the Projective Hypothesis, namely the premise that responses to projective tests would reveal 'the personality process or *private worlds* of individuals, ... that peculiar, individual way of organizing experience and of feeling' ([16, p. 392], italics in the original). The concept of projective anthropomorphism draws upon the broad conceptual underpinnings of projective techniques, not the methodology. Nevertheless, a closer look at the two most famous ones could be conducive to clarifying the present concept.

The Rorschach Inkblot test has been used in clinical settings since its publication by Rorschach in 1921. The procedure elicits responses to images presented in a standard series of ten inkblots. One of the most famous images resembles a winged creature, and most people readily 'see' a butterfly or a bat. Other inkblots are more ambiguous. Plate 2 could trigger an anthropomorphic reaction (the present

author sees two seated people with red turbans facing each other with hands touching), although [19] gives examples of widely diverse associations evoked by this inkblot. In terms of cognitive mechanisms, responses to inkblots fall into phenomena of pareidolia, i.e. the tendency to see meaningful patterns in a random stimulus such as clouds, rocks, cracks in the ground, and the Moon's surface [20]. Clinical practitioners who use the Rorschach or similar techniques, however, analyse patients' free associations towards uncovering their idiosyncratic thought patterns, defences and unconscious preoccupations. Turkle indirectly extrapolated the Projective Hypothesis to HRI when suggesting that 'robotic creatures become enhanced in their capacities to enact scenarios in which robots are Rorschachs, projective screens for individual concerns' [21, p. 2]. The concept proposed by the present author starts with a similar premise, but extricates the phenomenon in focus from the human–robot dyad (the next section expands).

The TAT was first developed by Henry Murray and associates in the 1930s in accordance with Murray's theory of psychogenic needs [22]. The procedure elicits imaginative stories of what is happening in a standard set of pictures printed on large cards, which show ambiguous situations. The original set includes also a blank card. The stories that are projected onto these cards are believed to disclose personal needs and motivations. It may be speculated that 'robot' (as an object constructed in discourse and the imagination) functions somewhat like a TAT picture insofar as we imaginatively embed it in storylines that anthropomorphise the machine.

5 Instances of projective anthropomorphism

As seen, psychological projection can imply functionally different processes in different contexts. The kind of projection that takes place during the performance of a projective test is arguably the closest to the kind of projection postulated here. However, although for some individuals robots might serve as 'projective screens' (to paraphrase [21]), the very existence of robots in modern society—the fact that our species creates them—begs the question, why? The purpose of articulating a concept of *projective anthropomorphism* in the first instance is to bring this question to the fore by labelling an aspect of robot-centred anthropomorphism that hitherto has been understated in the field.

In the immediate context, an apt analogy is the H–T–P test. To put it whimsically, it is as if on some collective level, our species was given the task of drawing Human, and has come up with a picture of humanlike automata. Fictive stories, legends and myths that humanity has long been telling about these projective creations reveal 'that

peculiarly human way of organising experience and of feeling’ (to paraphrase [16]). While projective techniques vary in their specific theoretical commitments, they all assert the creative aspect of the mind and emphasise the essentially holistic nature of mental processes [23]. It is proposed here that projective anthropomorphism rests on a creative process whereby symbolic representations of holistic experience of our own being are formed in imaginings of the humanoid automaton, and imbue this leitmotif with an aura of meaningfulness.

An additional reason for distancing projective anthropomorphism from Turkle’s characterisation of robots as ‘Rorschachs’ is that a robot is not a neutral image. Unlike an inkblot, it has a definite physical appearance and programmed mannerism that give the robot its character. Furthermore, a robot is not a passive ‘screen’ onto which anything could be projected. Robot companions or SARs are designed to be responsive. As Turkle herself pointed out, *Paro* makes eye contact, is sensitive to touch, and customises its behaviour according to whether it is being stroked gently or aggressively [21]. Moreover, robots come into persons’ lifeworld as artefacts already loaded with cultural associations and expectations, which reflect culturally specific traditions (e.g. in Japan as opposed to the USA; see a summary in [1]). This cross-cultural diversity could be explored in further elaboration of projective anthropomorphism (beyond this paper’s scope).

Phenomena of projective anthropomorphism include not only myths, legends, and science-fiction depictions of humanlike automata, but also linguistic tropes and deliberate ways of talking about actual robots. Technology has its own momentum, and as a consequence meanings of the word ‘robot’ continue to evolve. The software robot that we must prove we are not by checking a reCAPTCHA is far removed from the kind of robots that have populated modern imagination since Karel Čapek introduced this word to science fiction in his 1921 play *R.U.R.: Rossum’s Universal Robots*, and the robots imagined by Isaac Asimov, who coined the term ‘robotics’ in his fiction. Yet, tropes and images in popular culture, news media, and even academic papers, continue to perpetuate the ‘archetypal’ robot, the mechanical doppelgänger. For example, the humanoid form is impractical for robotic systems in many workplace applications, and yet there seems to be an irresistible urge to depict them as humanoid. In 2017, news media were full of doom and gloom headlines of how robots will take our jobs. In 2018, readers were given the good news: robots will create new jobs. Whether the narrative is dystopian or utopian, blog articles are often embellished with images of humanoid robots. Journalists deliberately choose such images because they resonate with readers through shared

cultural references; but the fact that the motif persists in popular culture begs a psychological explanation.

Similarly, the deliberate use of anthropomorphising language could be viewed as the ‘surface’ process that perpetuates a ‘deep’ process of meaning-making. For instance, a recent scholarly paper that comments on workplace automation from the standpoint of organization studies is subtitled ‘Why Robots Might Not Want to Steal Your Job’ [24]. To state the obvious, robots cannot really want anything. The author’s question is a rhetorical device, a humorous ‘hook’. In the humanities, anthropomorphism is usually defined as a literary device whereby writers ascribe humanlike traits, feelings, ambitions, or conduct to nonhumans (animals, natural phenomena, objects). Classic examples include Colodi’s *Pinocchio*, first published in 1883, which embodies the motif of a manmade humanoid automaton (*Pinocchio* is a wooden doll given the ability to talk, walk, think, and feel like a real boy). However, analyses of how tropes and other literary devices ‘work’ within a text towards conveying moral, political, or ideological messages, do not explain why this imago—the ‘archetypal’ humanoid robot—exists and persists.

6 Robot as Other

The ‘semiotic robot hypothesis’ [1] speculates that the leitmotif of humanoid automata fulfils some deep need in human beings, hence its staying power. The leitmotif survives because it is the fittest symbolic representation of some experiential state. To invoke Jung’s definition of *symbol*, ‘it is the best possible expression at the moment for a fact as yet unknown or only relatively known, ... standing for something that is only divined and not yet clearly conscious’ ([25] para. 817). Unlike an analogy or allegory, a symbolic representation in the Jungian sense is not a token or substitute for something known, and does not have a one-to-one correspondence to something else. Rather, it is a fantasy image, or imago, that draws together elements of bodily lived experiences into a gestalt-like whole that has its own meaning.

Occurrences of the robot leitmotif might be irreducible to any single psychodynamic principle. Nevertheless, it could be speculated that in general this imago relates to the fact that we experience ourselves as autonomous agents. As Fromm averred, ‘Man may be defined as the animal that can say “I”, that can be aware of himself as a separate entity’ [26, p. 60]. Being an animal that can see itself as a separate entity is associated also with a fear of death. Rank [27] suggested that mythological and literary motifs of the double might express a desire for immortality. To go a step beyond Rank, this may apply also to humanoid automata in myths,

folklore, and science fiction, and to some aspects of the discourse of social robotics.

Learning to see oneself as a separate entity is a developmental milestone in early childhood. Around eighteen months of age, typically developing toddlers start to recognise their mirror reflection. This can be verified in a procedure (traditionally called the rouge test) in which the experimenter discretely puts a smudge on the child's face, or a sticker on the child's head, without the child realising it, and then holds up a mirror for the child to see [28]. Children old enough to recognise themselves in the mirror reflection often touch their face or head when seeing the unexpected smudge or sticker. Some nonhuman animals also pass the mirror recognition test. Besides primates, the list includes elephants, dolphins, magpies, and more. However, human infants go through a phase of being fascinated with their own reflection. Lacan called it the 'mirror stage' [29]. According to his psychoanalytical theory, the external image produces in the young child a psychic response that gives rise to a mental representation of an 'I'. Lacan expanded his theory of the mirror stage in infancy to propose that human beings are driven by a sense of a fundamental lack of something inside us that could be called the self. Hence we always define ourselves in relation to external objects: to an Other, a not-me. In his theory, the mirror stage establishes the ego as dependent on the Other.

We are also the only animal that creates artefacts in its own image. The compulsion to create humanoid artefacts seems to have evolved with *homo sapiens*. Humanoid figurines carved from mammoth ivory were created between 40,000 and 35,000 years ago (the Upper Palaeolithic period in Europe). These include a lion-man sculpture discovered in Germany 1939, and numerous so-called Venus figurines discovered in various sites in Europe. Reportedly the oldest undisputed example of a depiction of a human being discovered to date is the Venus of Hohle Fels, found in 2008 in Germany. To archaeologists, Venus figurines serve as a source of information about prehistoric material culture [30]. A 1989 study that rigorously analysed the diversity of figurines found in European Russia concluded, 'We believe that each type of figurine had its own symbolic meaning' (quoted in [30, p. 515]). While the symbolic significance that prehistoric figurines had for their makers could never be known with any certainty, it is likely to have differed from the significance that robots have in modern society.

The psychological constant that is highlighted here is the fact that humans are motivated to take some material and convert it into something looking like us—and nowadays also something that behaves and talks like us. It is as if since prehistoric times our species has been compelled to bring forth humanoid artefacts so as to create mirror reflections of ourselves. It is as if wanting to understand how 'human' works, we try to build it.

7 Conclusion

The construct introduced here—projective anthropomorphism—is a means for defining a class of phenomena, potentially bringing its regularities to the fore, and thereby beginning to articulate its ensuing implications. As a conceptual tool, the 'proof of concept' lies in its timely capacity to open up spaces for exploring a dimension of the contemporary discourse of robotics that falls outside the remit of cognitive models of anthropomorphic reactions.

Phenomena of projective anthropomorphism can be found in social representations of robots in news media, documentaries, and other public domain contexts, which perpetuate themes and images of humanoids in popular culture, and may also precipitate motivations to build and invest in humanlike machines. While cultural heritages associated with attitudes to robots are widely acknowledged in the robotics-centred literature (see review in [1]), scholarly explorations are typically undertaken within diverse disciplinary contexts, each with its own priorities and webs of concepts. Conceptualising projective anthropomorphism could be a first step towards opening up issues that are not readily articulable within existing paradigms, and in this way expand the discourse of social robots beyond particular practical and ethical issues of immediate concern to robot designers and users. The utility of the novel concept and its implications for steering technological possibilities are yet to be explored in future work. At present, it is the author's epistemological position that identifying and analysing phenomena of projective anthropomorphism may help towards understanding how humanity's dialogue with itself is expressed in endeavours of social robotics.

Declaration

Conflict of interest The author declares that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Jones RA (2015) *Personhood and social robotics*. Routledge, London
2. Epley N, Waytz A, Cacioppo JT (2007) On seeing human: a three-factor theory of anthropomorphism. *Psychol Rev* 114:864–886
3. Zanatto D, Patacchiola M, Cangelosi A, Goslin J (2019) Generalisation of anthropomorphic stereotype. *Int J Soc Robot*. <https://doi.org/10.1007/s12369-019-00549-4>
4. Złotowski J, Sumioka H, Eyssel F, Nishio S, Bartneck C, Ishiguro H (2018) Model of dual anthropomorphism: the relationship between the media equation effect and implicit anthropomorphism. *Int J Soc Robot* 10:701–714
5. Turkle S (2011) *Alone together*. Basic Books, New York
6. Fong T, Nourbakhsh I, Dautenhahn K (2003) A survey of socially interactive robots. *Robot Auton Syst* 42:143–166
7. University of Hertfordshire (2019) Kaspar the social robot. www.herts.ac.uk/kaspar/the-social-robot. Accessed on 3 August 2019
8. Yao M (2017). Could your child's best friend be a robot? www.forbes.com/sites/mariyayao/2017/04/06/could-your-childs-best-friend-be-a-robot/#5c20b5a86dbb. Accessed 22 August 2017
9. Cog Project (n.d.) Overview. www.ai.mit.edu/projects/humanoid-robotics-group/cog/overview.html. Accessed on 2 August 2019
10. Brooks RA, Breazeal C, Marjanović M, Scassellati B, Williamson MM (1999) The cog project: building a humanoid robot. In: Nehaniv CL (eds) *Computation for metaphors, analogy, and agents CMAA 1998. Lecture notes in computer science*, vol 1562. Springer, Berlin
11. Brooks RA (2000) Will robots rise up and demand their rights? *Time* 155(25):86
12. Baumeister RF, Dale K, Sommer KL (1998) Freudian defense mechanisms and empirical findings in modern social psychology: reaction formation, projection, displacement, undoing, isolation, sublimation, and denial. *J Pers* 66:1081–1124
13. Sandler J, Perlow M (1989) Internalization and externalization. In: Sandler J (ed) *Projection, identification, projective identification*. Taylor and Francis, London, pp 1–12
14. Jung CG (1951) Aion. In: *The collected works of C G Jung*, vol 9 pt. II. Routledge and Kegan Paul, London
15. Sandler J (1989) The concept of projective identification. In: Sandler J (ed) *Projection, identification, projective identification*. Taylor and Francis, London, pp 13–26
16. Frank LK (1939) Projective methods for the study of personality. *J Psychol* 8:389–413
17. Gordon RM, Rudd-Barnard A, Smith-Wexler L (2018) House-tree-person test. In: Kreutzer JS, De Luca J, Caplan B (eds) *Encyclopedia of clinical neuropsychology*. Springer, Cham
18. Lilienfeld SO, Wood JM, Garb HN (2000) The scientific status of projective techniques. *Psychol Sci Publ Interest* 1:27–66
19. McCully RS (1987) *Jung and Rorschach*. Spring Publications, Dallas
20. Taylor RP, Martin TP, Montgomery RD, Smith JH, Micolich AP, Boydston C et al (2017) Seeing shapes in seemingly random spatial patterns: fractal analysis of Rorschach inkblots. *PLoS ONE* 12(2):e0171289. <https://doi.org/10.1371/journal.pone.0171289>
21. Turkle S (2006) New complicities for companionship: a nascent robotics culture. Talk Given at Stanford University, Stanford
22. Murray HA (1943) *Thematic apperception test manual*. Harvard University Press, Cambridge
23. Semeonoff B (1976) *Projective techniques*. Wiley, London
24. Fleming P (2019) Robots and organization studies: why robots might not want to steal your job. *Organ Stud* 40:23–37
25. Jung CG (1921) Psychological types. In: *Collected works of C. G. Jung*, vol 6. Routledge and Kegan Paul, London
26. Fromm E (1956) *The sane society*. Routledge, London
27. Rank O (1971) *The double*. University of North Carolina Press, Chapel Hill
28. Rochat P (2003) Five levels of self-awareness as they unfold early in life. *Conscious Cogn* 12:717–731
29. Lacan J (1977) *Ecrits: a selection*. Tavistock, London
30. Soffer O, Adovasio JM, Hyland DC (2000) The “Venus” figurines textiles, basketry, gender, and status in the upper Palaeolithic. *Curr Anthropol* 41:511–537

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Raya A. Jones, PhD, is a Reader at the School of Social Sciences, Cardiff University, where she teaches psychology. Her long-term research interests include Jungian, dialogical and narrative perspectives on the self, and more recently the discourse of social robotics. Her latest authored book is *Personhood and Social Robotics* (Routledge 2016). Earlier books include *Jung, Psychology, Postmodernity* (Routledge, 2007), *The Child-School Interface* (Cassell, 1995), and several edited and co-edited volumes, of which the latest is *Narratives of Individuation* (Routledge, 2019).